

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
29 December 2005 (29.12.2005)

PCT

(10) International Publication Number  
**WO 2005/124594 A1**

(51) International Patent Classification<sup>7</sup>: **G06F 17/30**,  
G01C 21/36, G09B 29/00

(21) International Application Number:  
PCT/IB2005/051924

(22) International Filing Date: 10 June 2005 (10.06.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/580,159 16 June 2004 (16.06.2004) US

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS, N.V.** [NL/NL];  
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(71) Applicant: **U.S. PHILIPS CORPORATION** [US/US];  
1251 Avenue of the Americas, New York, NY 10020 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DIEDERIKS, Elmo**

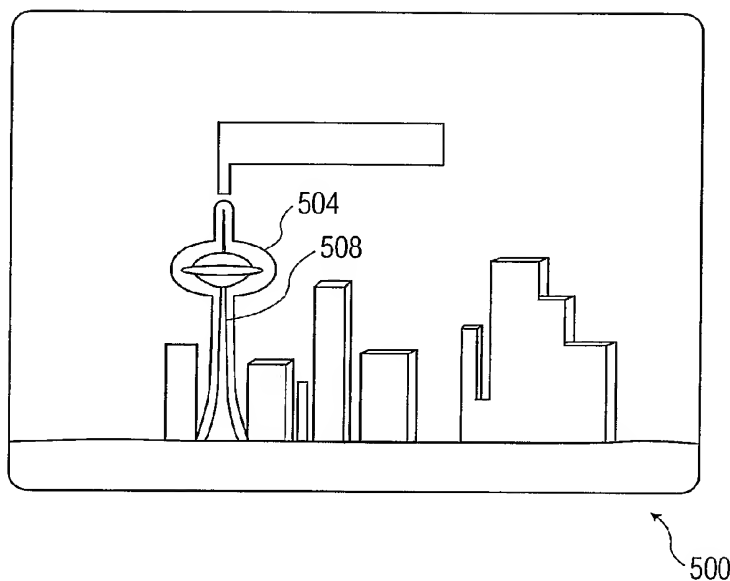
**Marcus Atilla** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). **VAN DER SLUIS, Bartel Marinus** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). **KORST, Johannes Henricus Maria** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). **CORTENRAAD, Hubertus Maria Rene** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(74) Common Representative: **KONINKLIJKE PHILIPS ELECTRONICS, N.V.**; c/o Russell Gross, P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GI, GM, GR, GU, HK, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: AUTOMATIC, REAL-TIME, SUPERIMPOSED LABELING OF POINTS AND OBJECTS OF INTEREST WITHIN A VIEW



(57) Abstract: A device operable with an image-viewing apparatus identifies known points/objects of interest in an image and superimposes onto the image being viewed labels to identify the known points/objects. Identification is made based on stored information for each point/object in a world model database and based on device location and orientation parameters that are updated automatically in real time.

WO 2005/124594 A1



(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL,

SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations

**Published:**

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

AUTOMATIC, REAL-TIME, SUPERIMPOSED LABELING OF POINTS AND OBJECTS  
OF INTEREST WITHIN A VIEW

5       The present invention relates to superimposing, onto an image, labels of points and objects appearing in the image. More particularly, the present invention relates to automatically in real-time determining if a predetermined point or object of interest appears in an image and, if so, superimposing the corresponding label onto the image.

10       It is known to determine location and orientation of a camera taking pictures. The registration of location for the purpose of easy browsing and retrieval of the pictures at a later time is described in U.S. Patent Nos. 5,506,644 to Suzuki et al. and 6,222,985 to Miyake (hereinafter "Miyake"), the disclosure of the latter being hereby incorporated by reference in its entirety. It is also known to add direction/orientation information to pictures. For instance, U.S. Patent No. 5,614,981 to Bryant et al. (hereinafter "Bryant"), the disclosure of which is hereby fully incorporated by reference herein, proposes a camera with a magnetic  
15       orientation sensor. U.S. Patent 5,659,805 to Furlani et al. describes a camera for indicating camera orientations on photographic film, and U.S. Patent No. 5,699,444 to Palm describes how image data (known objects on the photograph of which the exact location is known) can be used to determine the orientation of the camera. U.S. Patent No. 4,183,645 to Ohmura et al. relates to detecting the orientation of the photo (portrait or landscape) in order to add the  
20       date and time in the appropriate orientation. As a further application regarding orientation, U.S. Patent No. 6,304,284 to Dunton et al. (hereinafter "Dunton"), whose disclosure is herein incorporated by reference in its entirety, uses orientation information to create panoramic images with a camera.

25       Global positioning system (GPS) receivers which deliver real-time ground coordinate and altitude readings for the receiver can now be integrated into a single chip and are available to be built into all kinds of portable devices such as mobile phones and digital cameras, at low cost. There exists, furthermore, an observable trend of more and more digital photos and photo albums being shared on the Internet.

30       What is needed is a way to utilize readily accessible location and orientation data real-time to enhance the experience of taking a photo.

      The present invention has been made to address the above-noted shortcomings in the prior art. It is an object of the invention to provide a device, a method operable on the device, and a computer program for performing the method, wherein a device usable with an image-viewing apparatus marks, automatically in real-time, an image captured from a viewpoint.

The image is bounded in two dimensions by a field of view at a particular orientation. The device communicates with a storage medium for storing location information of a point of interest and/or an object of interest. The device includes a processing unit configured to, automatically in real-time, update a value of at least one parameter that defines the capturing  
5 of the image from the viewpoint. The processor also determines, based on the location information and at least one updated value, whether the respective point or object is disposed within the field of view at said viewpoint at said particular orientation. Further included in the device is a label superimposing unit configured to, automatically in real-time, based on the determination that the respective point or object is so disposed, and under control of the  
10 processing unit, superimpose a label onto the image to identify the respective point or object.

Details of the invention disclosed herein shall be described with the aid of the figures listed below, wherein:

FIG. 1 is a schematic diagram of a real-time label superimposition system in accordance with the present invention;

15 FIG. 2 is a how an image with superimposed labels appears to the user through the viewfinder of a camera in accordance with the present invention;

FIG. 3 is a conceptual diagram of a two-dimensional field of view in accordance with the present invention;

20 FIG. 4 is a flowchart of a process for superimposing a label for a point/object of interest within the viewed image in accordance with the present invention;

FIG. 5 is depiction of a sub-process of mapping a stored shape to the viewed image in accordance with the present invention;

FIG. 6 is a flowchart of a sub-process for refining label placement in accordance with the present invention using the sub-process depicted in FIG. 5;

25 FIG. 7 is a flowchart of a method for retrieving labels and associated metadata to be displayed upon playback or re-presentation of recorded images; and

FIG. 8 is a flowchart of a process for retaining label information for subsequent re-presentation of the stored images.

FIG. 1 shows, by way of illustrative and non-limitative example, a configuration for a  
30 real-time label superimposition system 100 according to the present invention. The real-time label superimposition system 100 can be realized, for example, as a modified form of the digital camera in FIG. 1 of the above-mentioned Miyake patent. The system 100 includes front-end optics 104, a view finder 108, a GPS unit 112, a liquid crystal display (LCD) monitor 116, an orientation sensor 120, a processing unit or processor 124, a user input device  
35 128 and a world model 132, all connected on a data and control bus 136.

As in Miyake, the front-end optics 104 may include a taking lens comprising a focus lens and a variable magnification lens, a diaphragm, and an optical low pass filter, these being arranged in series.

5 The view finder 108 has a front-end LCD 110, as in Miyake, but this LCD functions as a label superimposition unit. As such, the LCD 110 is preferably extended across the entire viewing lens of the view finder 108 to allow label superimposition anywhere in the image.

The GPS unit 112 is similar to that in Miyake, continuously providing updates as to ground coordinates and altitude.

10 LCD monitor 116 is optionally provided with the capability to display the superimposed labeling which has been embedded into the processing unit's video output to the LCD monitor.

The orientation sensor 120 detects camera tilt and horizontal rotation, as by the use of micro-machined silicon ("MEMS") sensors. Such sensors are discussed in the Dunton patent. One sensor is preferably utilized to detect tilt and another to detect rotation. Alternatively, the tilt sensor in the Bryant patent of a conventional film camera can be implemented. It is noted here that, in addition to orientation, the zooming/magnification factor of the variable magnification lens in the front-end optics 104 must be known to determine whether a point of interest or object of interest is currently within the field of view of the camera 100, as discussed below in further detail.

20 Processing unit 124 may include the Miyake strobe unit, charge-coupled device (CCD) unit, image pickup unit, camera unit, display unit, process unit and output unit. Accordingly, the image emitted from the low-pass filter in the front-end optics 104, as well as being displayed in the view finder 108 after label superimposition, is incident upon the CCD unit within the processing unit 124. The latter is necessary for storage of the image as a digital photo and for display on the LCD monitor 116.

25 The Miyake switch unit may be integrated into the user input device 128 of the present invention. Along with a release switch for taking photos, the switch unit includes a user-operable zooming/magnification factor selector. The user input device 128 preferably includes, in addition, a wheel, button or other mechanism for choosing among displayed labels in the image, and possibly other input means by which to filter, by classification, labels to be displayed. As an example of the latter, the user can filter tourist information (e.g. this is the Eiffel Tower), topographical information (e.g. Paris with 10,000,000 inhabitants, and so on). The user input device 128 may additionally feature a mechanism, such as a trackball, joystick or rocker switch, by which to navigate a cursor across the screen of the LCD monitor 35 116. This user interface can be utilized, for example, to limit concurrent display of labels to

points/objects of interest in the immediate vicinity of the cursor, so as to not overcrowd the image with labels.

World model 132 is a database, which may be incorporated into the Miyake PC card. The database 132 includes points of interest and objects of interest 140, and optionally shapes 144 of the objects. Loading of the database 132 ahead of time is especially useful if the camera is integrated with a personal digital assistant (PDA). The database 132 may alternatively be accessed via a wireless connection to the system or camera 100 from an Internet service affording WiFi (IEEE 802.11 wireless Ethernet) communication. Such access is especially useful if the camera 100 is integrated with a mobile phone. As a further alternative, this "travel-guide" information can be loaded into another portable device such as a PDA, which can wirelessly connect to the camera 100 via a Body Area Network or Personal Area Network.

Processing itself may also be off-loaded to a remote server (not shown) to reduce form factor and power requirements for the mobile image-viewing device. Updates from the GPS unit 112 and orientation sensor 120, for example, can be wirelessly transmitted to the server. The server determines whether objects/points of interest are within the field of view and, if so, where. The server transmits this determined data back to the camera 100 so that the processor 124 can appropriately operate the view finder 108 to superimpose labeling.

FIG. 2 portrays two exemplary views 204, 208 from a digital camera incorporating the real-time label superimposition system 100 of the present invention. The left-hand view 204 shows a label 212 having a stem portion 216 and containing a description 220 of an object of interest, namely the Matterhorn mountain. Other labels 224 are positioned near other respective objects or points of interest. An example of a point of interest is a particular location within a nature trail that affords a desirable view of a canyon. By manipulating the user input device 128, the user can switch to any of the other labels 224, thereby generating a descriptive label at the chosen location. Accordingly, the right-hand view 208 displays a label 228 that identifies Mont Blanc. In the case of the nature trail, for instance, the descriptive label may provide a number or other identifier that matches with a travel guide or hiking route. The auto-labeling function could be editable by the user by means of the user input device 128. Thus, the user could add information to labels or even create new labels such as "Home," "starting point of the trip," "my hotel room," and so forth. Non-descriptive labels only appear when more than one point/object of interest is concurrently displayed. Otherwise, if only a single point/object of interest is displayed, the only label appearing is the descriptive label. Alternatively, the user might override the default to allow more than one descriptive label, if applicable, to concurrently display.

The LCD 110 in the view finder 108, responsive to electrical signals, selectively obscures incident light to form the letters 232 "Mont Blanc" and the label outline 236. Optionally, the obscuring elements may be colored to form a colored label within the image 208. As an alternative to the use of LCDs, other light sources such as light-emitting diodes (LEDs) may be used in combination panels that are transmissive in one direction and reflective in the other direction to introduce into a common light path both the original image and the superimposed label.

FIG. 3 conceptually depicts an example of a field of view in two dimensions in accordance with the present invention. Since the image 302 as shown in FIG. 3 is rectangular, i.e., viewable through a rectangular viewing lens of the view finder 108, the field of view is defined simply by a horizontal sector angle 303 and a vertical sector angle 304, and, more generally, by the focal length of the taking lens in the front-end optics 104 and by the horizontal and vertical dimensions 306, 308 of the viewing lens of the view finder 108. Such truncation of the image 302 to the dimensions 306, 308 may occur anywhere in the light path, starting at the front-end and continuing through the view finder 108, and is shown in FIG. 3 at a particular location merely for simplicity of demonstration. In effect, the field of view is defined by exterior walls 312, 316, 320 of a solid figure, in this case a pyramid 324, whose apex coincides with the viewpoint 328, such that the walls extend through a periphery 332 of the image 302. The apex or viewpoint 328 is the optical center of the taking lens in the front-end optics 104. A focused rendition of the image formed at a focal length beyond the viewpoint 328 is routed to the view finder 108, where a label may be superimposed according to the invention. That rendition is also made incident upon a front surface of the CCD within the processing unit 124. The variable magnification or "zoom" lens in the front-end optics 104 features multiple different focal lengths, each entailing a different field of view. Although the field of view is defined in FIG. 3 by a pyramidal shape, implementation of the invention in, for example, binoculars would limit the field of view in accordance with a conical shape, e.g. to a single, uniform sector angle. The field of view, more generally, may be defined with respect to any arbitrary shape for the view finder 108.

Based on current ground coordinates and elevation provided by the GPS unit 112, tilt and rotational angle provided by the orientation sensor 120 and represented by the arrow 336, the processing unit 124 can determine by means of known geometrical transformations which, if any, point/objects of interest 140 within the world model 132 are currently within the current field of view. For example, as shown in FIG. 3, an object 328 at world coordinates (X,Y,Z) is currently within the field of view and displayed in the view finder 108 as object image 332 at image or view finder coordinates (x,y). Accordingly, the label (not shown)

stored in the world model 132 for object 328 is automatically placed in real-time adjacent to the object image 332 to identify the object.

Preferably, the world model 132 includes for the objects/points of interest 140 topographical data so that the system 100 accounts for the possibility of occlusion of the object/point behind an object in the foreground. Such modeling techniques are often utilized, for example, in on-screen portrayal of three-dimensional objects in computer graphics applications, such as computer animations. If the object/point is totally occluded, either the object/point is not labeled or, optionally, the stem on the label may be dotted to indicate occlusion.

FIG. 4 sets forth one example of a real-time label-superimposing process 400 for a point/object of interest within the viewed image in accordance with the present invention. The device of the present invention may be incorporated into a camera, camcorder, binoculars or night-viewing equipment, for example. If the image-viewing apparatus is on (step S404), the viewpoint and orientation are read. Both of these parameters may be repeatedly updated synchronously or asynchronously and buffered for reading. In addition, the most-recently buffered zooming/magnification factor can be read, to determine the most recent output of the user-operable zooming/magnification factor selector (step S408). For each point/object of interest in the world model 132 that has not already been linked to a label (as will be explained), a decision is made as to whether the point/object is within the current field of view (FOV) at the current viewpoint at the current orientation (step S412). This database search may, alternatively, be limited to points/objects within the FOV that are approximately at a particular distance from the viewpoint. Thus, as in Miyake, a focusing sensor is typically provided, within the camera unit of the processing unit 124, for calculating a distance from the camera 100 at which image focusing occurs. Labeling can accordingly be limited to only those objects/points within the FOV that are also at that distance or within an interval centered at that distance. In any case, the location within the image 302 is determined for each point/object within the FOV, optionally inclusive of merely those disposed approximately at the calculated focusing distance. This determined location is used to update the average location, if any yet exists, for that point/object. An average location is desirable, to reduce noise which appears as random error in the readings from the orientation sensor 120 and the GPS unit 112. The randomness of the errors tends to make the errors offset each other when the readings are averaged. The appropriate label is then superimposed onto the image 302 for viewing by the user (step S416). As more and more readings for a given object/point are averaged, jitter in the label is mitigated and stability is achieved. When people take pictures, for example, it can easily take 5-10 seconds to find the right composition



and settings for the camera, whereas multiple readings for an average can be accumulated quickly, e.g., within a fraction of a second. Query is made by the processing as to whether the determined location for the object/point, and correspondingly for the label, is stable, based for instance on a predetermined number of readings (step S420). If more readings are needed, the process repeats, starting at step S404. Preferably, the needed number of iterations is achieved for a single, point/object before processing a next point/object. If, on the other hand, stability has been achieved, a corresponding label is linked in fixed "on-screen" relation to tracked contours in the image 302 (step S424). Techniques in detecting contours in an image are well-known. Tracking can be implemented using the same methodology conventionally applied in motion compensation, examining candidate images within a predetermined window in a subsequent frame for the best-matching image. Once the label is linked to the tracked contours, location averaging is no longer needed. The label stems continues to point to the object/point even if the camera 100 is moved by the user, as long as the point/object remains within the field of view. A separate process (not shown) can monitor linked labels, and drop the link whenever the point/object exits the field of view.

FIG. 4 represents merely an example of the specific design. Thus, for example, the query as to location stability (step S420) may, in addition to determining whether the average has been computed with the predetermined number of summand, also take into account detection by the MEMS sensors of recent significant camera movement.

The repeated updating of parameters, as a further example, may include retention of respective rolling averages for the viewpoint and orientation, the rolling average having a set number of summands. The query as to location stability in step S420 would then be reduced to identifying when stability of the rolling average had not yet been achieved, e.g., when the camera has just been turned on or when the MEMS sensors detect recent significant movement in the camera. In addition, the query as to existence of non-linked points/objects with the FOV in step S412 may, in a single iteration, determine all such points/objects. Likewise, superimposition of labels in step S416, and label linking in step S424, would be performed in a single iteration.

FIG. 5 illustrates a stored-shape mapping sub-process 500 for mapping to the viewed image a shape 504 stored among the object shapes 144 in the world model 132 in accordance with the present invention. The sub-process 500 reduces systematic error in the orientation and GPS readings, and may be executed between steps S420 and S424 of the real-time label-superimposing process 400. The shape 504 of the object of interest 508, in this case the "Toronto Tower," is, as shown in FIG. 6, first mapped to the average location determined in step S416 once the location has been determined to be stable in step S420. Then, a process

similar to that described above for contour tracking may be performed. Examination is made within a search window of the mapped image for a best match, the search window being of predetermined size (step S608). The average location is then updated by the translation determined to align the mapped image with the best matching candidate image (step S612).

- 5 The corresponding label placed to point to the updated location is then linked in step S424, thereby reducing systematic error.

A preferred embodiment not only minimizes random and systematic error in the parameter readings, but reduces the storage/processing overhead by limiting the world model database 132 to data specific to the objects/points of interest. In particular, the database 132  
10 need not include data throughout three-dimensional space to enable determination as to whether occlusion of the point/object exists. It suffices to model merely the points/objects of interest to include characteristics such as, for example, shape, color and shading, these being stored for multiple possible orientations or alternatively in real-time extracted from a three-dimensional model. This expands upon the concept introduced in FIG. 6 of storing shape  
15 information. As in the FIG. 6 embodiment, once the object/point is determined to be within the FOV, known image recognition techniques can be employed to decide whether the object/point is sufficiently occluded so as to warrant dotted labeling or no labeling at all, depending on the implementation. More specifically, after execution of the location averaging sub-step (in step S416), query is made as to whether the point/object is visible in  
20 the image 302. This query involves retrieving the stored characteristics for the point/object and comparing them within a window of the image 302, that window being centered at the average location in a manner similar to that of step S608. The comparison may involve, for instance, merely a contour based on the retrieved characteristics as matched against a detected contour within the window. It is also noted that the image characteristics to be matched need  
25 not be exactly co-located with the point/object, but may, instead, lie within a vicinity of the point/object.

In any event, if a predetermined level of pixel-to-pixel matching is not attained, the object/point is judged to be not visible in the image. Otherwise, if the match is close enough, the average location is updated as in step S612, not only reducing systematic error, but, in  
30 addition, accounting for occlusion or non-occlusion of the object/point. This approach may be further modified, as mentioned above, by retention of respective rolling averages for the viewpoint and orientation.

Despite paring down the world model 132 by the above-described method, the extra storage and processing overhead for the characteristic data may suggest off-loading, from a  
35 mobile device, the world model and much of the image processing to the aforementioned

server. With respect to the image processing, the camera/binoculars 100 may transmit the averaged parameter readings to the server (step S408). The server determines potential points/objects of interest that are included within the FOV, not taking into account occlusion. For each point/object, it determines how the object would look in the image 302, based on the received parameter readings. At least partial images are transmitted back including approximately where in the image 302 they are located. The above-described pixel-to-pixel matching is then applied. To prevent a wireless-bandwidth bottleneck in the event of multiple currently-viewed points/objects of interest, the images transmitted back can be limited to particular camera-requested categories.

Video captured by a camcorder, for example, can be downloaded to a personal computer (PC) that efficiently identifies points/objects of interest while playing back the video. As shown in FIG. 7, the image-viewing apparatus, such as a camcorder, retains, in storage, contours that are linked to a label in step S424 (step S704). The stored video and contours are downloaded to the PC (step S708), or may be physically transferred as by means of a PC card. As the video plays back, the contour(s) are matched to the playback image to identify the corresponding points/objects of interest. Such identification also makes available for presentation any metadata that has been stored in connection with the contour, such as the descriptive label or an identification based on GPS and a mapping database as to the town or other location of the user at the time the video was recorded. Although a delay may exist between the time a contour or set of contours first appears on the video and the time at which the real-time label superimposition system 100 "locks" onto the contour(s) in step S424, the matching by the PC on playback identifies the object/point as soon as the appropriate contour(s) appear. Advantageously, the labeling in the video playback may therefore first occur for an image recorded earlier than the image for which the labeling first appeared in the view finder of the camcorder.

The contour retention and matching also applies to the digital photos stored for a digital camera such as that described with respect to FIG. 1. As a photo is taken, the contour(s) for visible points/objects of interest, and associated metadata, are retained for subsequent matching of the contours to photos displayed on the PC. Once a match is made, the metadata is displayed on-screen, superimposed on or alongside the respective photo. In this way, vacation photos are much easier to relive.

Referring to FIG. 8, even the need to rematch contours can be eliminated, by storing the updated label location in the frame every time the video or photo frame is stored. In particular, since the label is linked, with respect to location, in fixed relation to the currently tracked contour(s) (step S424), and since the image coordinates of the currently tracked

contour(s) are known, the image coordinates of the label are known and this information, along with any description, may be embedded in the frame to be stored. When the stored photos or video are replayed on the PC, these stored coordinates are read and used to correctly superimpose the stored labels. Simultaneously, any other metadata stored with the frame can be accessed for visual display, aural presentation as by means of a speech synthesizer, etc.

5 Query is first made as to whether a video or photo frame is to be stored (step S804). If so, the label location which may have been updated (step S612) and is linked in fixed relation to the contour(s) (step S424) is stored along with the frame to be stored. In particular, an identifier of the label location, and any metadata, may be stored as part of the frame or in  
10 linked relation to the frame (steps S808, S812). During re-presentation or playback of the frame on the PC, for example, the PC superimposes the label based on the identifier and performs any further presentation based on the metadata. This playback technique may also be applied, for example, to superimpose labeling in the LCD monitor 116 concurrently with image/video capture, in which case, the frame, label location(s) and description are stored in a  
15 buffer.

As has been demonstrated above, a user-friendly adjunct to a digital camera, binoculars or other image-viewing apparatus automatically provides in real time labels to identify known points/objects of interest. An enhanced image-viewing experience is afforded, particularly for the tourist.

20 While there have been shown and described what are considered to be preferred embodiments of the invention, it will, of course, be understood that various modifications and changes in form or detail could readily be made without departing from the spirit of the invention. For example, the user may be provided with an actuator on the user input device 128 by which to introduce onto or exclude from the stored digital photos labels displayed in  
25 the view finder. It is therefore intended that the invention be not limited to the exact forms described and illustrated, but should be constructed to cover all modifications that may fall within the scope of the appended claims.

## WHAT IS CLAIMED IS:

1. A device (100) usable with an image-viewing apparatus (S404), said device for marking, automatically in real-time, an image (302) captured from a viewpoint (328), said image being bounded in two dimensions by a field of view (303, 304) at a particular orientation (336), said device in communicative connection with a storage medium (132) for storing location information of at least one of a point of interest and an object of interest (140), said device comprising:

a processor (124) configured to, automatically in real-time, update (S408) a value of at least one parameter that defines the capturing of said image from said viewpoint and determine, based on said location information and one or more of the at least one updated value, whether the respective point or object is disposed within said field of view at said viewpoint at said particular orientation (S412); and

a label superimposing unit (110) configured to, automatically in real-time, based on the determination that the respective point or object is so disposed, and under control of said processor, superimpose upon said image a label to identify said respective point or object (S416).

2. The device of claim 1, wherein said one or more includes a value of at least one of said viewpoint and said particular orientation (S408).

3. The device of claim 1, wherein said one or more includes a value of a zooming/magnification factor (S408).

4. The device of claim 1, wherein said field of view is defined by exterior walls (312, 316, 320) of a solid figure (324) whose apex coincides with said viewpoint, said walls extending through a periphery (332) of said image.

5. The device of claim 1, wherein the superimposing includes:

calculating where, in said image, said respective point or object is located based on said location information and one or more of the at least one updated value (S416); and  
arranging the label in said image based on the calculated location (S416).

6. The device of claim 5, wherein said updating is performed iteratively in real-time to generate a plurality of values for viewpoints, orientations or both (S408), so that the superimposing is performed based on an average location for the label (S416) and entails tracking movement of image content and linking said average location to the tracked content such that the label remains disposed in fixed relation to said content during the movement of said content (S424).

7. The device of claim 6, wherein the superimposing entails detecting a contour (S424), said image content comprising said contour.

8. The device of claim 1, wherein the storing step stores an object of interest (140).

9. The device of claim 8, wherein said location information includes a shape (144) of the stored object of interest and the superimposing entails:

calculating where in said image said object is located (S416); and

arranging the label in said image based on the calculated location (S416), said

arranging involving:

based on the calculated location, mapping the stored shape to said image (S604); and

aligning the mapped shape to a shape in the image within a predetermined distance of the mapped shape (S608).

10. The device of claim 1, wherein the storing stores a plurality of points and/or objects of interest, the determining and superimposing being performed automatically in real-time for the plural points and/or objects (S412).

11. The device of claim 10, further including a user input device (128) for selecting among labels of respective ones of said plurality superimposed in said image for concurrent presentation (224), the currently selected label (220) presenting information descriptive of the respective point or object.

12. The device of claim 1, further including said image-viewing apparatus into which said device is integrated, said image-viewing apparatus including at least one of a camera and a pair of binoculars (S404).

13. A system comprising the integrated device of claim 12, said system further comprising said storage medium for storing location information for at least one of a point of interest and an object of interest, wherein said processor is further configured to read the stored location information and to control the label superimposing unit based on said information (112, 120, S412, S416), said device being further configured to present for viewing by a user said image containing the superimposed label (108, 116).

14. The device of claim 1, wherein the label has a portion for pointing to said respective point or object (216).

15. A method for marking an image automatically in real-time, said method comprising the steps of:

storing, in a storage medium, location information for at least one of a point of interest and an object of interest (140); and

providing an apparatus (100) that automatically in real-time performs the steps of:

capturing (104), from a viewpoint (328), an image (302) bounded in two dimensions by a field of view (303, 304) at a particular orientation (336);

updating a value of at least one parameter that defines the capturing of said image (S408);

determining, based on said location information and one or more of the at least one updated value, whether the respective point or object is disposed within said field of view at said viewpoint at said particular orientation (S412); and

based on the determination that said respective point or object is so disposed, superimposing upon said image a label to identify said respective point or object (S416).

16. The method of claim 15, wherein said one or more includes a value of at least one of said viewpoint, said particular orientation and a zooming/magnification factor (S408).

17. The method of claim 15, wherein the superimposing includes the steps of:

calculating where in said image said respective point or object is located based on said location information and one or more of the at least one updated value (S416); and

arranging the label in said image based on the calculated location (S416).

18. The method of claim 17, wherein the identifying is performed iteratively in real-time to generate a plurality of viewpoints, orientations or both (S412), so that the

superimposing step comprises the step of deriving an average location for the label (S416),  
the superimposing step further comprising the steps of:

tracking movement of image content (S424); and

linking said average location to the tracked content such that the label remains  
disposed in fixed relation to said content during the movement of said content (S424).

19. The method of claim 15, wherein the storing step stores an object of interest, said  
location information includes a shape (144) of the stored object of interest and the  
superimposing step comprises the steps of:

calculating where in said image said object is located (S416); and

arranging the label in said image based on the calculated location, said arranging step  
further including the steps of:

based on the calculated location, mapping the stored shape to said image (S604); and

aligning the mapped shape to a shape in the image within a predetermined distance of  
the mapped shape (S608).

20. A computer-readable medium for storing a program (124) that includes  
instructions executable by a processor in performing automatically in real-time the steps of:

capturing (104), from a viewpoint (328), an image (302) bounded in two dimensions  
by a field of view (303, 304) at a particular orientation (336);

updating a value of at least one parameter that defines the capturing of said image  
(S408);

determining, based on location information for at least one of a point of interest and  
an object of interest and based on one or more of the at least one updated value, whether the  
respective point or object is disposed within said field of view at said viewpoint at said  
particular orientation (S412); and

based on the determination that said respective point or object is so disposed,  
superimposing upon said image a label to identify said respective point or object (S416).



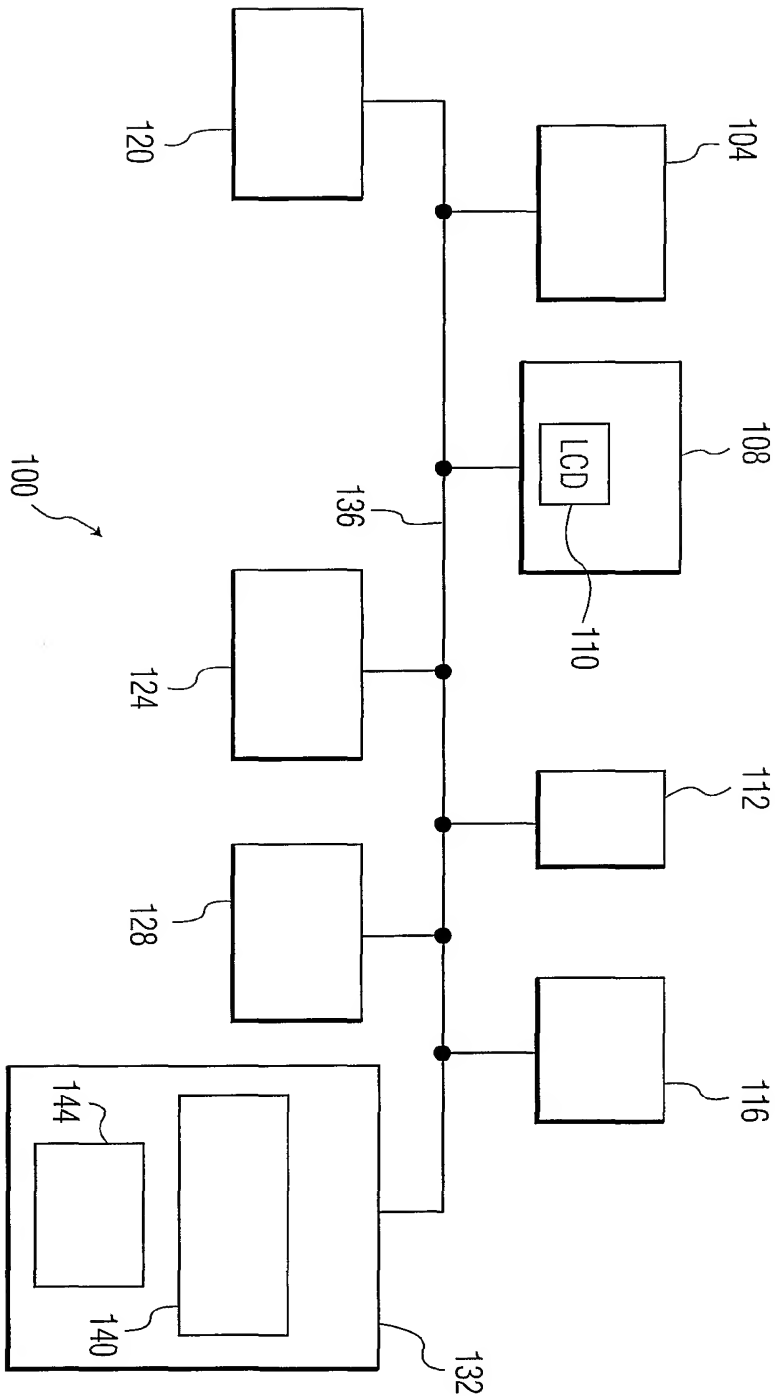
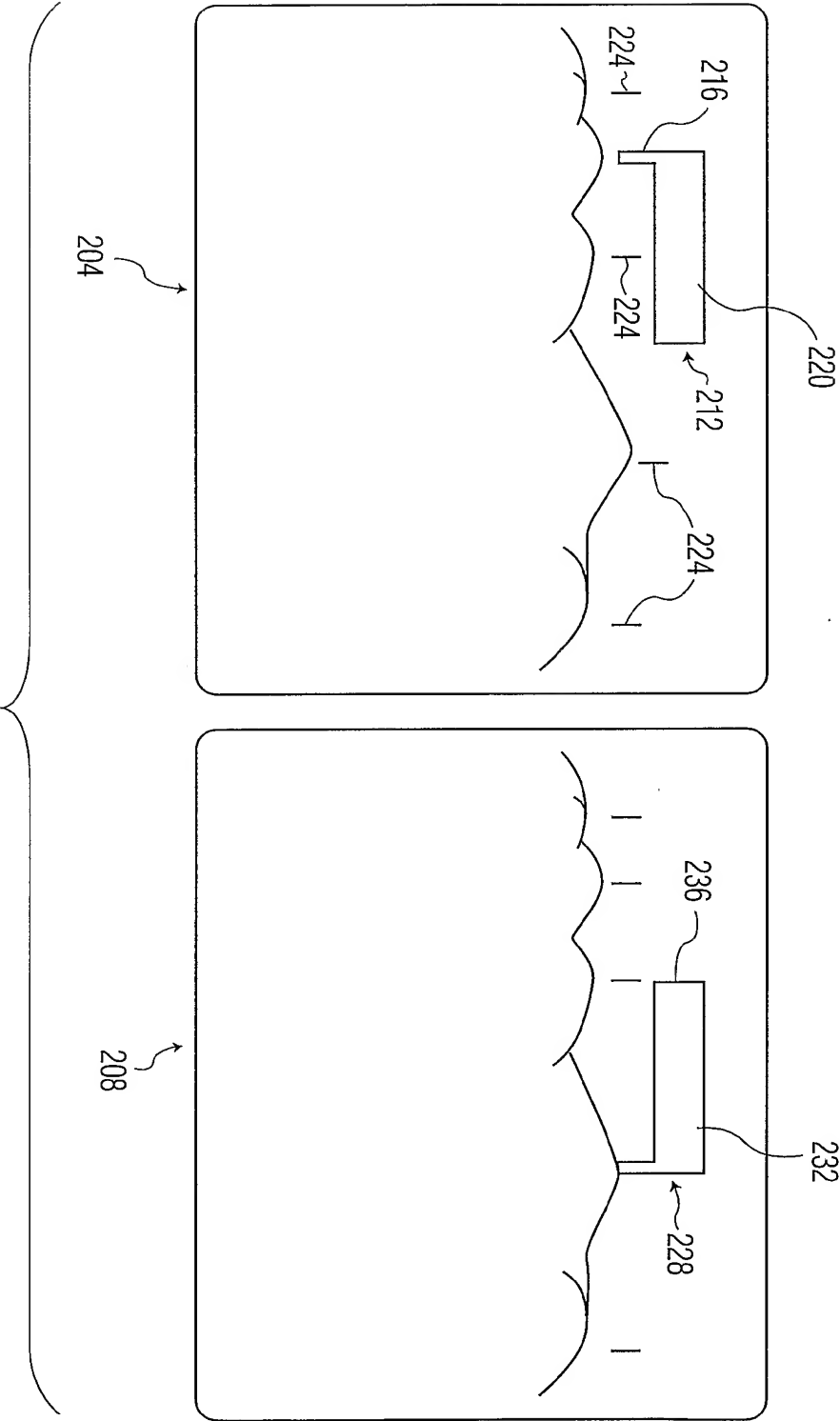


FIG. 1



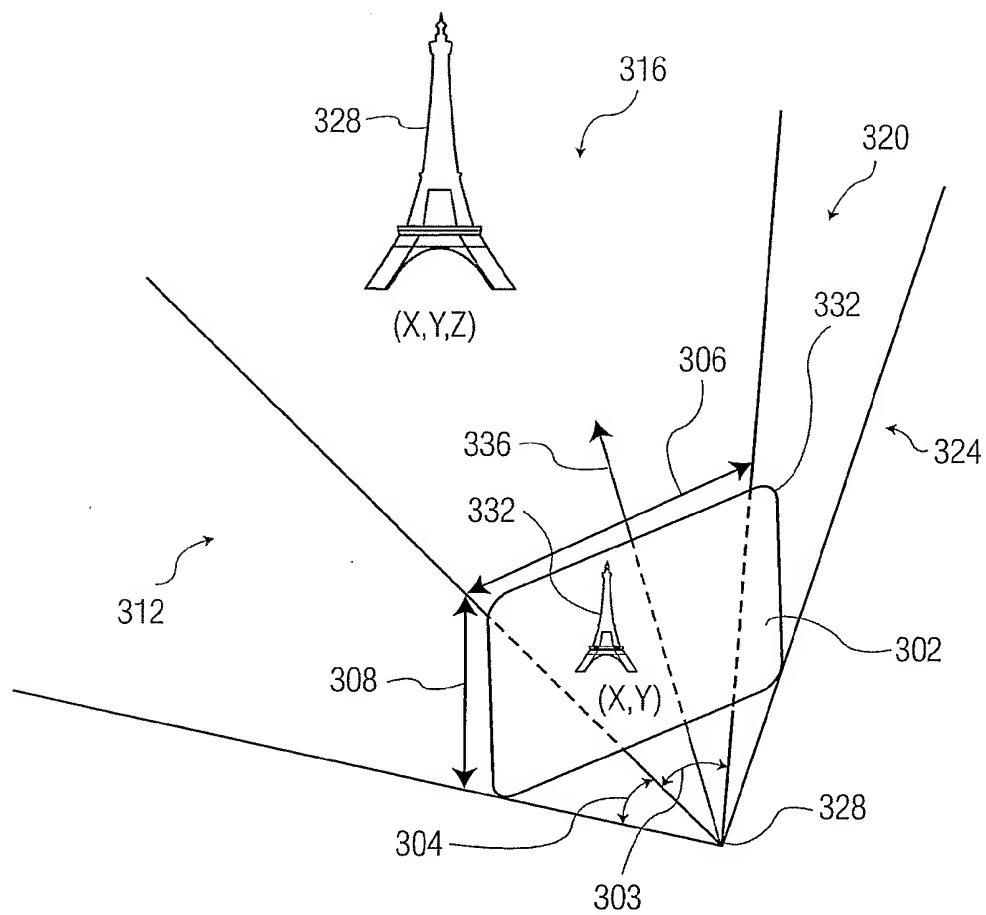


FIG. 3

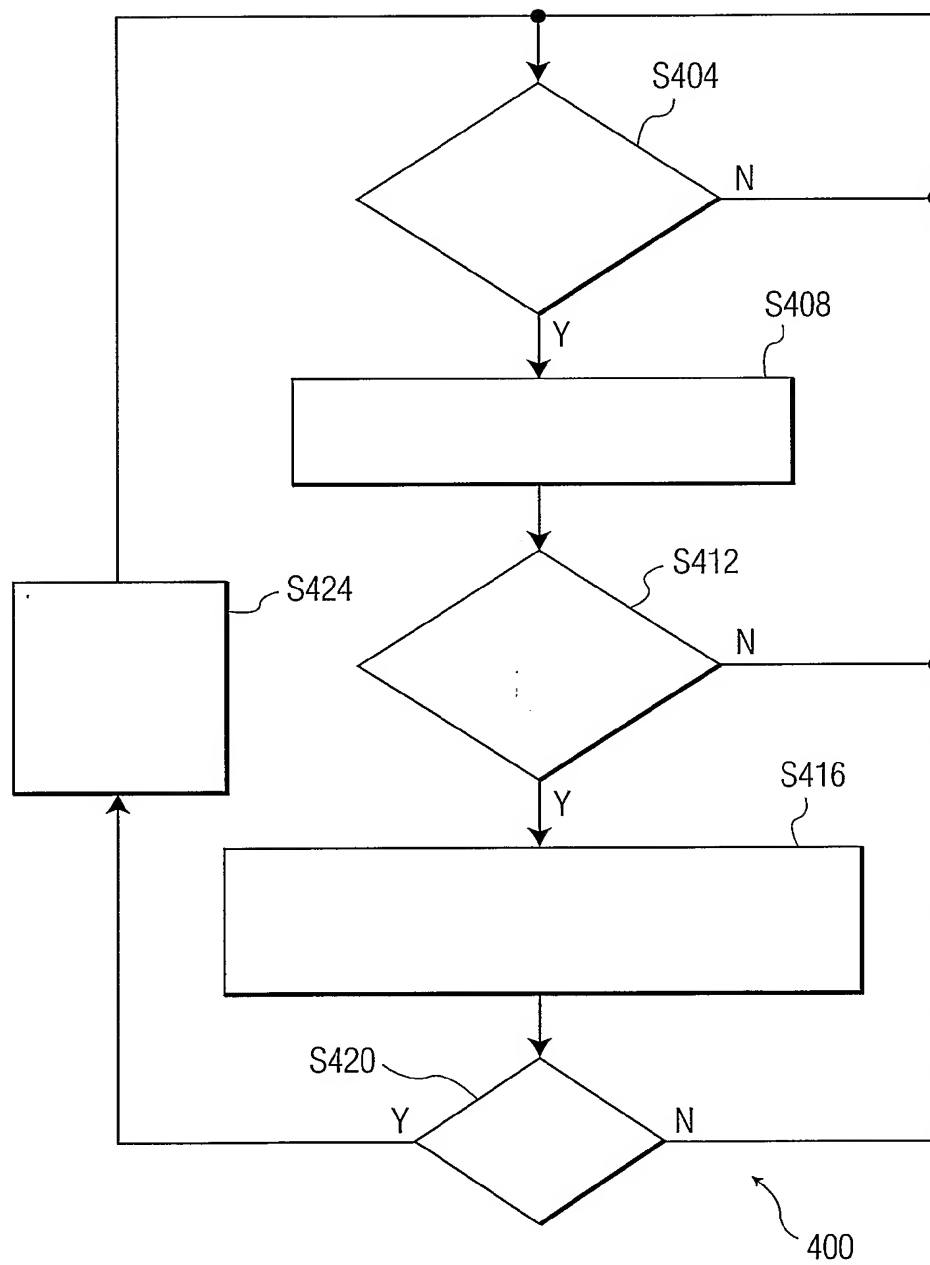


FIG. 4

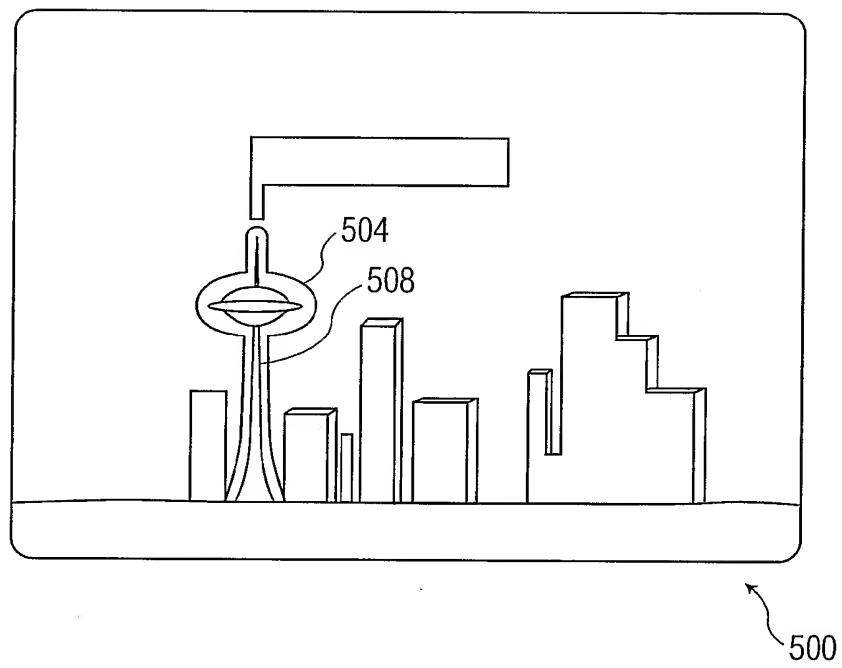


FIG. 5

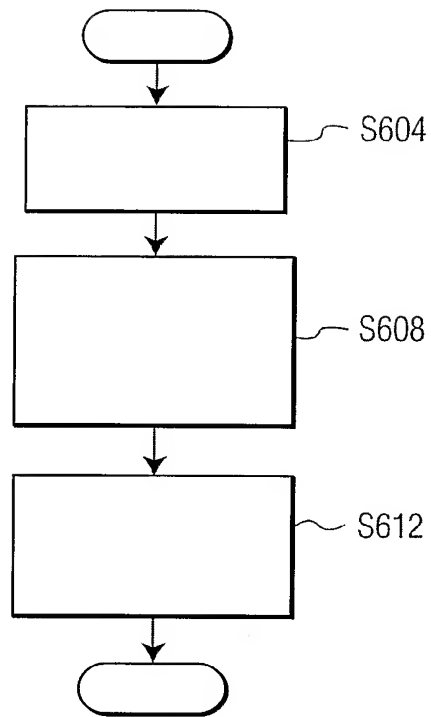


FIG. 6

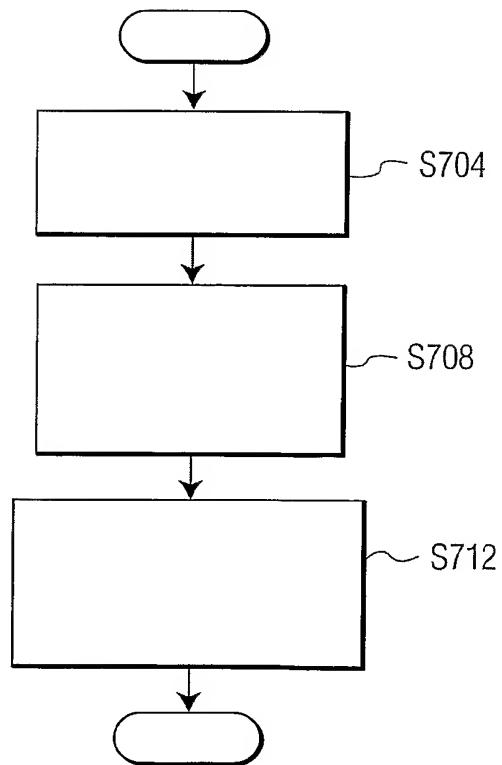


FIG. 7

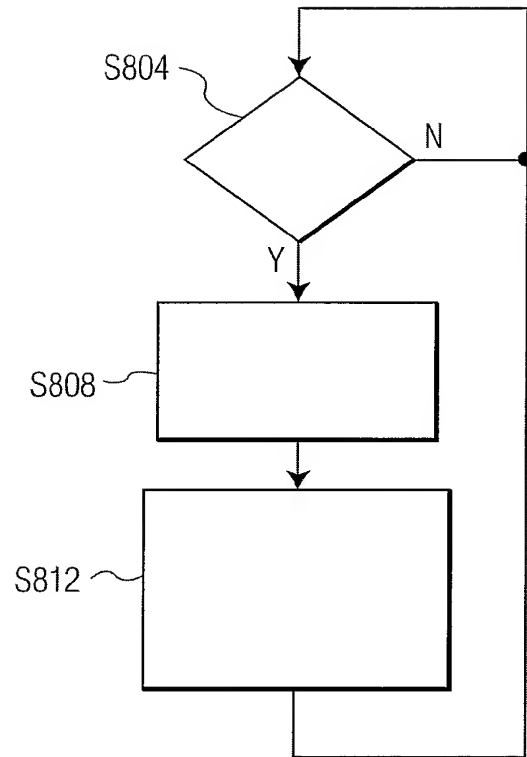


FIG. 8

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IB2005/051924

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06F17/30 G01C21/36 G09B29/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G01C G06T G09B H04N G06F G09G

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, COMPENDEX, INSPEC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2002/075282 A1 (VETTERLI MARTIN ET AL) 20 June 2002 (2002-06-20)	1-6,8, 10-18,20
Y	the whole document	7,9,19
X	US 2002/052684 A1 (BIDE STEPHEN) 2 May 2002 (2002-05-02)	1-6,8, 10-18,20
Y	the whole document	7,9,19
X	US 6 282 362 B1 (MURPHY MICHAEL ET AL) 28 August 2001 (2001-08-28)	1-6,8, 10-18,20
Y	abstract figures 1,2	7,9,19
A	US 6 222 985 B1 (MIYAKE IZUMI) 24 April 2001 (2001-04-24) cited in the application abstract	1-6,8, 10-18,20
	----- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

5 September 2005

Date of mailing of the international search report

19/09/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

PAVON



## INTERNATIONAL SEARCH REPORT

Int. Patent Application No.  
PCT/JP2005/051924

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 01/01075 A (BETHERE; KOROCHKIN, CARL, P) 4 January 2001 (2001-01-04) abstract figure 48	7,9,19
A	----- US 5 699 444 A (PALM ET AL) 16 December 1997 (1997-12-16) cited in the application abstract -----	7,9,19

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IB2005/051924

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002075282 A1	20-06-2002	US 6208353 B1 EP 1246080 A2	27-03-2001 02-10-2002
US 2002052684 A1	02-05-2002	AT 231236 T AU 739611 B2 AU 7779198 A CA 2307129 A1 CN 1262734 A DE 69810768 D1 EP 0986735 A1 GB 2325975 A ,B WO 9855833 A1 JP 2002502499 T	15-02-2003 18-10-2001 21-12-1998 10-12-1998 09-08-2000 20-02-2003 22-03-2000 09-12-1998 10-12-1998 22-01-2002
US 6282362 B1	28-08-2001	NONE	
US 6222985 B1	24-04-2001	JP 10210336 A JP 10210337 A JP 10271375 A CN 1193116 A ,C CN 1574905 A CN 1574906 A CN 1574907 A CN 1574908 A EP 1411721 A1 EP 1427204 A1 EP 1411722 A1 EP 0855835 A2 US 2001010549 A1	07-08-1998 07-08-1998 09-10-1998 16-09-1998 02-02-2005 02-02-2005 02-02-2005 02-02-2005 21-04-2004 09-06-2004 21-04-2004 29-07-1998 02-08-2001
WO 0101075 A	04-01-2001	US 6912293 B1 AU 5635100 A WO 0101075 A2	28-06-2005 31-01-2001 04-01-2001
US 5699444 A	16-12-1997	NONE	